

# Bayesian Network:

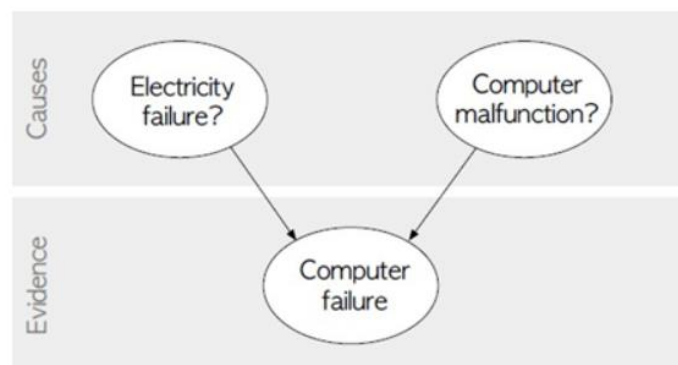
A Bayesian network is a directed acyclic graph in which each edge corresponds to a conditional dependency, and each node corresponds to a unique random variable.

Bayesian network consists of two major parts: a directed acyclic graph and a set of conditional probability distributions

- The directed acyclic graph is a set of random variables represented by nodes.
- The conditional probability distribution of a node (random variable) is defined for every possible outcome of the preceding causal node(s).

For illustration, consider the following example. Suppose we attempt to turn on our [computer](#), but the [computer](#) does not start (observation/evidence). We would like to know which of the possible causes of computer failure is more likely. In this simplified illustration, we assume only two possible causes of this misfortune: electricity failure and computer malfunction.

The corresponding directed acyclic graph is depicted in below figure.



The goal is to calculate the posterior conditional probability distribution of each of the possible unobserved causes given the observed evidence, i.e.  $P[\text{Cause} \mid \text{Evidence}]$ .

## Data Set:

**Title:** Heart Disease Databases

The Cleveland database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that has been used by [ML](#) researchers to this date. The "Heartdisease" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4.

# Attribute Information:

1. age
  2. sex
  3. chest pain type (4 values)
  4. resting blood pressure
  5. serum cholestoral in mg/dl
  6. fasting blood sugar > 120 mg/dl
  7. resting electrocardiographic results (values 0,1,2)
  8. maximum heart rate achieved
  9. exercise induced angina
  10. oldpeak = ST depression induced by exercise relative to rest
  11. the slope of the peak exercise ST segment
  12. number of major vessels (0-3) colored by flourosopy
  13. thal: 0 = normal; 1 = fixed defect; 2 = reversable defect
- The names and social security numbers of the patients were recently removed from the database, replaced with dummy values.